Thinking Like a Skeptic: Defeasible Inference in Natural Language

Year: 2020 | Citations: 111 | Authors: Rachel Rudinger, Vered Shwartz, Jena D. Hwang, Chandra Bhagavatula, Maxwell Forbes

---

Abstract

Defeasible inference is a mode of reasoning in which an inference (X is a bird, therefore X flies) may be weakened or overturned in light of new evidence (X is a penguin). Though long recognized in classical AI and philosophy, defeasible inference has not been extensively studied in the context of contemporary data-driven research on natural language inference and commonsense reasoning. We introduce Defeasible NLI (abbreviated \delta-NLI), a dataset for defeasible inference in natural language. Defeasible NLI contains extensions to three existing inference datasets covering diverse modes of reasoning: common sense, natural language inference, and social norms. From Defeasible NLI, we develop both a classification and generation task for defeasible inference, and demonstrate that the generation task is much more challenging. Despite lagging human performance, however, generative models trained on this data are capable of writing sentences that weaken or strengthen a specified inference up to 68% of the time.